DNA sequences of group 4 allergens from rye, wheat, barley and *Lolium perenne*

Comparison with isoforms of *Phleum pratense* Phl p 4

A. Nandy, M. Wald, L. Gräfe, O. Cromwell, H. Fiebig

Allergopharma Joachim Ganzer KG, R&D Department, 21465 Reinbek, Germany
Contact e-mail: andreas.nandy@allergopharma.de

Introduction

Grass pollen allergy is one of the most important allergic diseases world-wide. Several grass species grown in meadows, like *P. pratense* and *L. perenne*, contribute to allergic sensitisations, but also allergens from extensively cultured cereals, especially rye, make a profound contribution to the development of allergy. The group 4 major allergen of *P. pratense*, Phl p 4, is recognised by more than 70 % of grass allergic patients[1,2,5]. IgE-binding cross-reactivity has been described for some group 4 allergens of different grass species[3], but until now only the Phl p 4 gene could be deciphered on the DNA-level.

Results

The Pooideae group 4 allergens represent a family of basic proteins with molecular weights of about 55 kDa and calculated pI values far above 8 (Tab. 1, Fig. 1). In rye, wheat and *P. pratense* distinct isoforms with amino acid identities of 88 to 94 % could be detected. Additionally these isoforms exist in different minor variants. The inter-species homology lies in the range 83 % (Phl p 4 to Triticeae species) to 95 % (Sec c 4 to Tri a 4) (Fig. 2, Fig. 3).

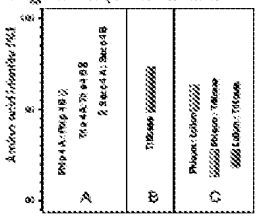
Tab. 1 Sequence analysis of grass pollen group 4 allergens

Protein	Source	Sequence length (amino acids)	Isoelectric point (pI)	Molecular weight (Da)
Phl p 4 A	*Phleum pratense*	509	8,8	55.895
Phl p 4 B	*Phleum pratense*	509	9,2	55.924
Lol p 4*	*Lolium perenne*	423 (fragment)	8,8*	*
Sec c 4 A	*Secale cereale* (rye)	486	9,1	54.939
Sec c 4 B	*Secale cereale* (rye)	496	9,3	54.993
Tri a 4 A	*Triticum aestivum* (wheat)	497	8,8	55.237
Tri a 4 B	*Triticum aestivum* (wheat)	497	8,8	55.149
Hor v 4	*Hordeum vulgare* (barley)	498	9,3	54.815

The sequence length, isoelectric points and molecular weight calculations were made on the basis of the mature proteins. For Phl p 4 the N-terminal residue has been determined by N-terminal protein sequencing. Based on this homology alignment (Fig. 1) the putative cleavage sites of Triticeae spp. have been used for calculation.

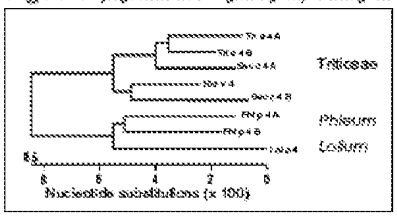
* The Lol p 4 sequence is only partial and contains about 85 % of the mature Lol p 4 sequence.

Fig. 2 Sequence identities



A: Intra-species identities of intra-species variants of group 4 allergens range from 89 % (Sec c 4 B) to 95 % (Tri a 4). The two major variants of Phl p 4 show inter-variant identities of 92 %.
B: The identities of allergens of the Triticeae species range from 96 % (Sec c 4 B to Tri a 4) to 95 % (Sec c 4 A to Tri a 4 B).
C: Inter-species identities of members of the genera *Phleum*, *Lolium* and of the 2 Triticeae genera *Hordeum*, *Triticum* and *Secale*.

Fig. 3 Phylogenetic tree of grass group 4 allergens

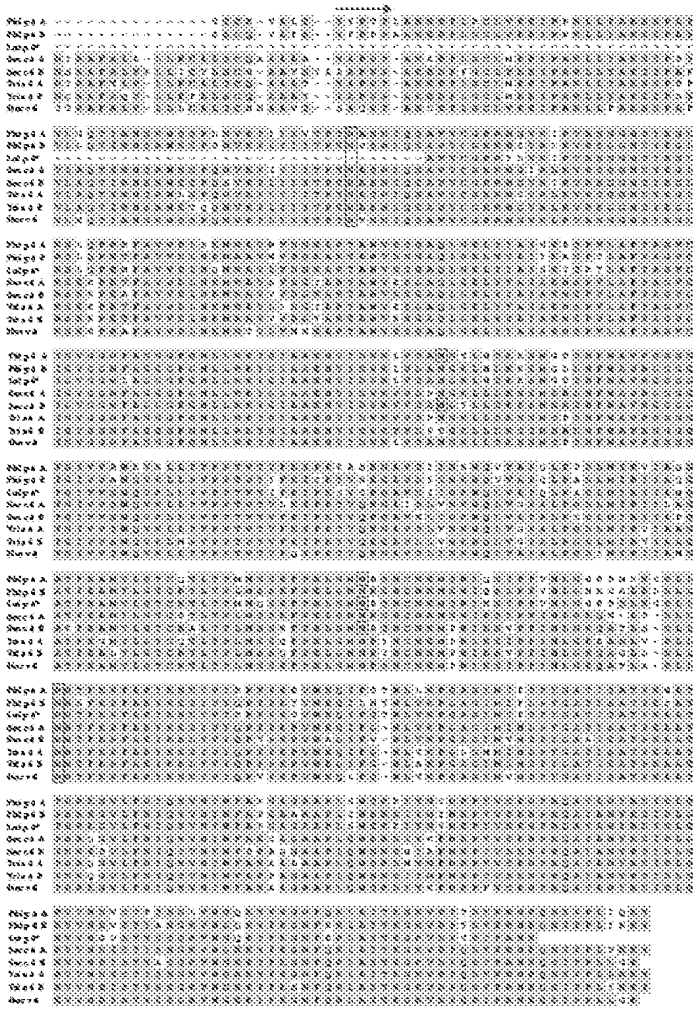


The cladogram* illustrates the phylogenetic relationships of the grass group 4 sequences. The rooted tree has been generated by using the DNA sub-sequences that overlap the Lol p 4 fragment (127 bp).
Remarkably intra-species variants (eg. Sec c 4 A and Sec c 4 B) show sequence identities similar to those of sequences originating from different Triticeae species (compare also Fig. 2A and 2B). The same can be seen for *Phleum pratense* variants that have similar degrees of amino acid difference as compared to the *Lolium perenne* sequence (compare also Fig. 2A and 2C).

Methods

Based on the DNA sequence of Phl p 4 several PCR-primer sequen with cross-reactivity to DNA sequences of related species could designed. The group 4 DNA sequences of *Lolium perenne* (Lol p *Secale cereale* (Sec c 4), *Hordeum vulgare* (Hor v 4), and *Triti aestivum* (Tri a 4) have been amplified, cloned and sequenced.

Fig. 1 Deduced amino acid sequence alignment of grass pollen group 4 allergens



Multiple alignment of *Phleum pratense* Phl p 4 variant forms, *Lolium perenne* Lol p 4, *Secale cereale* (rye) Sec c 4 variant forms, *Triticum aestivum* (wheat) Tri a 4 variant forms, and *Hordeum vulgare* (barley) Hor v 4. Residues that match the consensus sequence are yellow. The start of the mature Phl p 4 sequence as deduced by N-terminal protein sequencing of purified natural Phl p 4 is marked. Potential N-glycosylation sites are marked with blue boxes.

* The Lol p 4 sequence is only partial and contains about 85 % of the mature Lol p 4 sequence.

Conclusion

The group 4 allergens represent a family of proteins that are conserved among different grass species. The occurrence of cross-reacting isoforms in distinct species with amino acid homologies that are comparable to those of different group 4 molecules across the species border is remarkable. Since recombinant group 4 allergens may be important for a future

References

1) R.E. Rossi et al. (2001), Allergy 56, 1150-11853
2) K. Andersson and J. Lidholm (2003), Int. Arch. Allerg Immunol., Review article, 130, 87-107
3) S. Stumvoll et al. (2002), Biol. Chem. 383, 1983-19
4) Lasergene DNASTAR, Inc., Madison, WI 53715, U.S